

物化视图在大型 ORACLE 数据库查询中的应用

陈贻鑫, 倪子伟

(厦门大学 信息科学与技术学院 福建 厦门 361005)

【摘 要】: 随着网络的普及和电子商务的不断发展,以及数据库规模的不断扩大、信息量不断增加,对数据库的性能要求也不断提高,物化视图可以针对某些特定的常用查询,预先计算并保存表连接或聚集等花费较长时间的操作的结果,从而大大减少数据库查询统计的响应时间。

【关键词】: 大型 oracle 数据库;SQL 查询;物化视图;性能优化

1、引言

随着计算机的普及和信息网络用户爆炸性地增长,网络信息量呈几何增长,人们对数据信息的需求直接反应在对数据库系统的查询,但是对于数据量巨大的数据库而言,在进行复杂的查询时往往无法及时响应,等待时间超出用户所能接受的范围,不能满足应用要求,这就要求我们对数据库系统进行合理设置,从不同层面运用各种技术,提高查询的速度。

数据库优化技术有很多种,对于大型 Oracle 数据库而言,投影、选择、连接等操作需要耗费大量的时间,通过物化视图技术解决这类问题是一种很好的选择。

2、物化视图的概念

2.1 简介

Oracle 的物化视图可以看成是 SQL 查询的实例化,它在很多方面和索引有相似之处,最明显的相似之处就是两者都能提高数据库的查询性能,但它和索引的实现机制不同,它的优化原理是预先计算并保存表连接或聚集等耗时较多的操作的结果,即生成该查询结果某个时刻的快照,通过使用物化视图,诸如联接和聚合这些耗时较长的操作就不需要再重新执行了。它存储视图的定义以及执行视图所得到的行,和视图一样,它是以一个查询作为基础。然而,执行查询的同时创建视图,并将结果存储在表中。这样,在执行查询时,就可以避免进行这些耗时的操作,而从快速地得到结果。

当物化视图中的数据可以满足查询时,Oracle 服务器切换到引用物化视图,而不是基础表进行查询。

另外也可以将这张表定义和其他表具有相同的存储参数,可以将它放在您所选的表空间当中,还可以如同创建其他表的索引一样创建物化视图的索引以进一步改善对它们执行查询的性能。

2.2 创建语法

图一中:

[ENABLE|DISABLE QUERY REWRITE] 提供是否支持查询重写的选择;

[<REFRESH{FAST|COMPLETE|FORCE}][{ON COMMIT|ON DEMAND}] [NEVER REFRESH]指定数据刷新机制,如果需要进行快速刷新,则需要建立物化视图日志,增加 WITH 子句,且需要创建物化视图日志物化视图日志根据不同物化视图的快速刷新的需要,可以建立为 ROWID 或 PRIMARY KEY 类型的。还可以选择是否包括 SEQUENCE、INCLUDING NEW VALUES 以及指定列的列表。可以指明 ON PREBUILT TABLE 语句将物化视图建立在一个已经存在的表上。这种情况下,物化视图和表必须同名;

Creating Materialized Views: Syntax Options

```
CREATE MATERIALIZED VIEW mview_name
  [TABLESPACE ts_name]
  [PARALLEL (DEGREE n)]
  [BUILD {IMMEDIATE|DEFERRED}]
  [{ REFRESH {FAST|COMPLETE|FORCE}
  | [ON COMMIT|ON DEMAND]}
  | NEVER REFRESH ]
  [{ENABLE|DISABLE} QUERY REWRITE]
AS SELECT ... FROM ...
```

图一

[ON PREBUILT TABLE]将物化视图建立在一个已经存在的表上,这种物化视图的查询重写要求参数 QUERY_REWRITE_INTEGRITY 必须设置为 trusted 或者 stale_tolerated。

需要注意的是,查询并不总是执行创建物化视图后立即执行,它取决于生成 BUILD 子句。如果选择 IMMEDIATE 方式才是创建视图时同时建立查询的数据,而选择 DEFERRED 方式查询数据第一次刷新时建立。

2.3 物化视图和视图的区别

· 物化视图预先计算并保存表连接或聚集等耗时较多的操作的结果,这样在执行查询时,就可以避免进行这些耗时的操作,而从快速地得到结果;

普通视图在查询时,需要在查询中嵌套个子查询然后去查询原表;

· 使用物化视图的目的是为了提高查询性能,而普

通视图性能较低;

- 物化视图对应用透明,增加和删除物化视图不会影响应用程序中 SQL 语句的正确性和有效性;而普通视图会影响原表数据;

- 物化视图需要占用存储空间;而普通视图不需要;

- 当基表发生变化时,物化视图也应当刷新;而普通视图不需要。

3、物化视图创建的思路

- ①分析现有 SQL 语句,统计总结 SQL 的共同规律,作为设计物化视图的基本思路;

- ②定义相应物化视图,包括 SQL 查询语句、刷新方式的设置等;

- ③评价原有统计运算 SQL 语句的执行计划,是否被 Oracle 查询重写到相应物化视图;

- ④在物化视图上创建索引进行进一步优化;

- ⑤评估所有物化视图数据和索引的空间消耗情况,从而确定物化视图数据和索引表空间的容量。

4、物化视图的优势

4.1 不选用物化视图的弊端

如果不使用物化视图,那么你就必须根据常用投影、选择、连接等查询创建概要表以应对这些查询对大型数据库耗费大量的时间的操作,这样一来就需要花费大量的时间手动来创建它们、为它们建索引、为它们进行更新,并引导用户去使用它们。

然而在你花费了大量的时间创建出概要表之后,虽然从概要表执行 SQL 查询所需的时间的确会比原表少,但另一方面,你却还必须手动去维护更新概要表以保持它的数据和原表一致。

可以预见,这样的组织形式很快会在维护方面造成困难。

4.2 使用物化视图所能带来的好处

使用物化视图,数据库管理员就无需再手动创建概要,终端用户也不再需要去注意已定义的概要表,你可以创建一个或多个等价于概要表的物化视图。

以创建物化视图取代概要表的好处不仅在于物化视图将查询的结果“物化”到数据库的表中,而且还会生成查询重写引擎的元数据信息,用于自动重写 SQL

查询。一旦进行查询重写,Oracle 就会自动将能够通过查询物化视图得到结果的基表查询切换到物化视图来查询,即每当应用程序执行 SQL 查询,Oracle 服务器自动改写它转而使用相应物化视图。这样一来就避免了聚集、连接等操作,直接从已经计算好的物化视图中查询数据,极大减少了大型数据库查询的响应时间。

和创建概要表的方法相比,查询所需的响应时间相同,其主要的区别在于该应用程序不需要重写,重写进程是由系统自动处理的。

再者,物化视图提供了另一种重要的可选功能:自动刷新数据,物化视图根据用户预先设置的刷新模式刷新数据以保持和基表同步,这就免去了概要表需要数据库管理员手动维护更新同步的弊端,大大减轻了数据库管理员的工作负担,也减少了不必要的安全隐患。

5、结 论

物化视图是 Oracle 将查询重写功能和快照机制以及优化器结合起来创建的优化工具,它只需要扫描少量数据,减少物理读操作;同时它不需要经常排序聚集,从而减少了写操作;再者它已做好聚集计算和函数调用,查询时就不需要重复进行此类操作,减少 CPU 的消耗,由此使得用户访问大型数据库的查询速度可以增加好几倍,甚至可以用秒级的时间查询 TB 级的数据。

同时它也有不足之处,譬如物化视图本身需要占有一定的磁盘空间、物化视图进行数据同步刷新时也会对系统造成一定负担,但是相比起物化视图对大型数据库查询性能所带来的良好支持而言,这些缺点还不足以成为人们拒绝使用物化视图的理由。

参考文献:

- [1]Priya Vennapusa . Oracle Database 10g SQL Tuning Edition 2.0 D17265GC20[DB/OL].
<http://www.itpub.net/thread-522252-1-1.html>,2006-12.
- [2]James Spiller ,Jean-Francois Verrier . Oracle Database 10g Performance Tuning Edition 1.1 D19165CN11[DB/OL].
<http://www.itpub.net/thread-522252-1-1.html>,2009-11.
- [3](美)Loney.k 著,张立浩,尹志军译,Oracle Database 10g 完全参考手册.清华大学出版社,2006-9.

(上接第 164 页)

容的学习存在一定的困难,因此会借助于其它的一些学习平台解决问题。这也要求我们的精神课程能针对学生的具体问题,提供更好的学习资源,这也是精品课程所应努力的方向。

参考文献:

- [1] 郑庆华.Web 知识挖掘:理论、方法与应用[M].北京:科学出版社,2010.
- [2] <http://www.ncfuke.com>